

82478-1800
Hiroyuki Norishita et al.
JWP/949.253.2754

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 2 年 1 1 月 1 1 日
Date of Application:

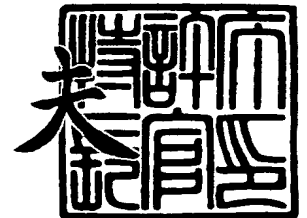
出 願 番 号 特 願 2 0 0 2 - 3 2 7 4 2 9
Application Number:
[ST. 10/C] : [J P 2 0 0 2 - 3 2 7 4 2 9]

出 願 人 松 下 電 器 産 業 株 式 会 社
Applicant(s):

2 0 0 3 年 9 月 1 6 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 3 - 3 0 7 5 8 5 7

【書類名】 特許願

【整理番号】 2022540460

【提出日】 平成14年11月11日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 12/08

【発明者】

 【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

 【氏名】 森下 広之

【発明者】

 【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

 【氏名】 清原 督三

【特許出願人】

 【識別番号】 000005821

 【氏名又は名称】 松下電器産業株式会社

【代理人】

 【識別番号】 100090446

 【弁理士】

 【氏名又は名称】 中島 司朗

【手数料の表示】

 【予納台帳番号】 014823

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9003742

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 キャッシュメモリ制御装置

【特許請求の範囲】

【請求項 1】 マイクロプロセッサ、キャッシュメモリ及びメインメモリを含み、複数のタスク処理を行うコンピュータシステムにおいて、キャッシュメモリを制御するキャッシュメモリ制御装置であって、

前記マイクロプロセッサから、実行中のタスクを識別するタスク ID を受け付けるタスク ID 受付手段と、

前記マイクロプロセッサから、メインメモリ上のアドレスを受け付けるアドレス受付手段と、

前記複数のタスクそれぞれに前記キャッシュメモリ内の別個のキャッシュメモリ領域を割り当てている割当手段と、

前記アドレス受付手段により受け付けたアドレスのデータが前記キャッシュメモリに格納されていない場合、前記メインメモリから当該アドレスのデータを含むデータ群を取得し、当該データ群を前記タスク ID 受付手段により受け付けたタスク ID と対応するキャッシュメモリ領域に格納する格納手段とを備える

ことを特徴とするキャッシュメモリ制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、キャッシュメモリを制御するキャッシュメモリ制御技術に関する。

【0002】

【従来の技術】

マイクロプロセッサとメインメモリとの中間に配置されるキャッシュメモリは、マイクロプロセッサがメインメモリ上のあるアドレス（以下、「メインメモリアドレス」という。）のデータにアクセスする際の、アクセス速度の遅さを緩和することを目的として用いられている、高速アクセス可能なメモリである。具体的には、マイクロプロセッサがあるメインメモリアドレスのデータにアクセスする際に、キャッシュメモリに当該データを含むデータ群をライン単位で一時的に

格納して、以後、マイクロプロセッサがメインメモリにアクセスしなくても高速アクセスが可能なキャッシュメモリにアクセスしてデータの読み書きが行えるようにしている。

【0 0 0 3】

ライン単位とは、キャッシュメモリに格納するデータサイズの一定容量を示す単位をいう。例えば、1 ラインを 1 6 バイトとすると、マイクロプロセッサから要求されたデータを含むデータ群を 1 6 バイト分読み出して、キャッシュメモリに格納する。このように、ライン単位でメインメモリ上のデータ群を一時的にキャッシュメモリに格納することをキャッシングという。

【0 0 0 4】

プログラムのメモリアクセスパターンには、アクセスされたメインメモリアドレスの近傍アドレスが、近い将来にアクセスされる確率が高いことを示す空間的局所性と、アクセスされたメインメモリアドレスが再び、近い将来にアクセスされる確率が高いことを示す時間的局所性の 2 つの性質があるので、ライン単位でデータ群をメインメモリからキャッシングしておけば、ある程度のキャッシュヒット率が期待できる。しかし、キャッシュメモリとして一般的に用いられる S R A M (S t a t i c R a n d o m A c c e s s M e m o r y) は高価であるため小容量であり、メインメモリ上の一部のデータしかキャッシングしておくことができない。そこで、キャッシュメモリを効率的に利用するために、メインメモリのデータをキャッシュメモリに格納する方法として、3 つのマッピング方式 (ダイレクトマッピング方式、セットアソシエィテブ方式、フルアソシエィテブ方式) が考案されている。

【0 0 0 5】

その中でもダイレクトマッピング方式は、マイクロプロセッサからデータアクセスがあった際に、キャッシュメモリ上のどのアドレス (以下、「キャッシュ内アドレス」という。) に該当データが格納されているかがすぐにわかり、探索に時間がかからないというメリットがある (非特許文献参照)。

【0 0 0 6】

【非特許文献】

斎藤忠夫・大森健児共著「現代 計算機アーキテクチャ原理」オーム社出版、
平成6年4月10日、p112

【0007】

【発明が解決しようとする課題】

しかしながら、ダイレクトマッピング方式では、データ群を格納するキャッシュ内アドレスはメインメモリアドレスの一部を用いて固定的に選定されるので、キャッシュメモリ内に空きエントリがあっても、選定されたキャッシュ内アドレスに既に格納されている有用なデータ群を捨てて、新たにキャッシングしたデータ群を書き込む場合があり、キャッシュメモリの効果を損なうという欠点がある。同様の欠点は、セットアソシエティブ方式においても起こる。

【0008】

特に、シングルマイクロプロセッサによる複数のタスクを擬似的に並行実行するマルチタスク処理や、複数のマイクロプロセッサ又はマルチプロセッサによる複数の異なるタスク処理において、各タスクの実行によってアクセスされるデータのアドレスは局所的でないので、いわゆるキャッシュミスとなり、キャッシングが頻発することになる。それに伴い、あるタスクの実行による新たなキャッシングにより、既に他のタスクの実行によってキャッシングされたデータ群が書き替えられる事象が頻発する。

【0009】

キャッシュメモリ内のデータ書き替えが頻発すると、データアクセスに費やされる時間は増加するので、結果的には各タスクのリアルタイム性の保証が難しくなる。

そこで、本発明は、シングルマイクロプロセッサによるマルチタスク処理や複数のマイクロプロセッサ又はマルチプロセッサによる複数のタスク処理において、各タスクのリアルタイム性を保証するべく、あるタスク実行による新たなキャッシングにより、他のタスク実行によってキャッシングされたデータ群が書き替えられないようにするキャッシュメモリ制御装置を提供することを目的とする。

【0010】

【課題を解決するための手段】

上記目的を達成するために本発明に係るキャッシュメモリ制御装置は、マイクロプロセッサ、キャッシュメモリ及びメインメモリを含み、複数のタスク処理を行うコンピュータシステムにおいて、キャッシュメモリを制御するキャッシュメモリ制御装置であって、前記マイクロプロセッサから、実行中のタスクを識別するタスク ID を受け付けるタスク ID 受付手段と、前記マイクロプロセッサから、メインメモリ上のアドレスを受け付けるアドレス受付手段と、前記複数のタスクそれぞれに前記キャッシュメモリ内の別個のキャッシュメモリ領域を割り当てている割当手段と、前記アドレス受付手段により受け付けたアドレスのデータが前記キャッシュメモリに格納されていない場合、前記メインメモリから当該アドレスのデータを含むデータ群を取得し、当該データ群を前記タスク ID 受付手段により受け付けたタスク ID と対応するキャッシュメモリ領域に格納する格納手段とを備えることを特徴とする。

【 0 0 1 1 】

【発明の実施の形態】

以下、本発明に係るキャッシュメモリ制御装置について、図面を用いて説明する。

<構成>

図 1 は、本発明に係るキャッシュメモリ制御装置を含むコンピュータシステムの主要構成を示す図である。本システムは、キャッシュメモリ制御装置 1、キャッシュメモリ 2、マイクロプロセッサ 3、主記憶制御装置 4、メインメモリ 5 等から成る。

【 0 0 1 2 】

マイクロプロセッサ 3 は、プログラムの実行や各種の演算などを行なう装置であり、CPU (Central Processing Unit) とも呼ばれる。本実施の形態において、マイクロプロセッサ 3 はマルチタスク処理を行うが、各タスクの数は予め決まっており、各タスクのワーキングセットも予めわかっているものとする。また、OS (Operating System) 制御下において、マイクロプロセッサ 3 は、実行するタスクが切り替るとキャッシュメモリ制御装置 1 に対してタスク ID を生成して送出する。タスク ID は、プログラ

ムが格納されているメインメモリアドレスを変換して生成してもよいし、当該メインメモリアドレスそのものであってもよく、タスクそれぞれを識別することが可能であればどのようなものであってもよい。

【0013】

キャッシュメモリ 2 は、メインメモリ 5 より高速にデータアクセスが可能な記憶媒体であり、例えば、SRAM である。また、メインメモリ 5 は、キャッシュメモリ 2 よりアクセス速度が低速な記憶媒体であり、例えば、DRAM (Dynamic Random Access Memory) である。

主記憶制御装置 4 は、メインメモリ 5 に対するデータアクセスをコントロールする装置であり、いわゆる、メモリコントローラである。

【0014】

キャッシュメモリ制御装置 1 は、一般的にはキャッシュコントローラと呼ばれ、キャッシュメモリ 2 に対するデータアクセスをコントロールする装置である。本実施の形態においてキャッシュメモリ制御装置 1 は、1 ラインにつき 16 バイト分のデータをキャッシュメモリ 2 に格納する。そして、ライン単位でキャッシュメモリ 2 に格納したデータ群を、キャッシュディレクトリと呼ばれるアドレス帳で管理している。キャッシュディレクトリは、ライン単位のデータ群を格納する、キャッシュ内アドレスとタグとを対応付けたテーブルである。キャッシュメモリ制御装置 1 は、ライン単位のデータ群をキャッシュメモリ 2 に格納する場合、当該データ群のメインメモリ上の位置を識別するためのタグを生成して、当該タグとキャッシュ内アドレスとを対応付けてキャッシュディレクトリに記憶する。

【0015】

また、キャッシュメモリ制御装置 1 は、マイクロプロセッサ 3 により実行されるタスクの数に応じてキャッシュメモリを領域分割して管理している。各領域と各タスクとの対応付け及び領域のサイズは、マイクロプロセッサ 3 からの設定命令（プログラム）により、各タスク実行前に設定される。

図 2 は、キャッシュメモリ制御装置 1 のブロック構成図である。キャッシュメモリ制御装置 1 は、タスク ID 入力レジスタ 11、領域指定部 12、領域管理部

13、アクセスアドレス分解部14、キャッシュ内アドレス生成部15、hit／miss判定部16等から成る。

【0016】

タスクID入力レジスタ11は、マイクロプロセッサ3から送出されたタスクIDを受け付ける。受け付けられたタスクIDは、領域指定部12に送られる。

領域指定部12は、マイクロプロセッサ3からの命令により設定された各領域とタスクIDとの対応を示す領域情報を保持している。領域情報には各領域と対応するタスクのタスクIDが記載されている。例えば、図2において、領域0とタスクAとが対応付けて設定されているので、領域0情報にタスクAのタスクIDが記載されている。領域指定部12は、各領域情報に記載されているタスクIDと、タスクID入力レジスタ11から送られてきたタスクIDとを、比較器(cmp)を介して比較して、タスクIDが一致した比較器のビットを立てて、領域を指定する領域番号を生成する。領域指定部12により生成された領域番号は領域管理部13に送られる。

【0017】

領域管理部13は、図3に示す領域管理表を保持している。領域管理表は、領域番号、ベースアドレス及びサイズを一組として対応付けた表であり、マイクロプロセッサ3から受けた設定命令により設定される。領域番号は、割り当てられた各領域を識別する番号であり、ベースアドレスは、各領域の開始位置を示すキャッシュ内アドレスである。また、サイズは、各領域の容量を示す値であり、値1が1ライン分の容量に相当する。

【0018】

領域管理部13は、領域指定部12から送られてきた領域番号を受けて、領域管理表を参照し、当該領域番号と対応するベースアドレスをキャッシュ内アドレス生成部16に、サイズをアクセスアドレス分解部14に送る。

アクセスアドレス分解部14は、マイクロプロセッサ3から送出されたメインメモリアドレスを受け付けて、当該メインメモリアドレスを3つのフィールドに分解して、タグ、インデックス、バイトセレクトとする。タグは、受け付けたメインメモリアドレスのデータを含むライン単位のデータ群を識別するための符号

として用いられる。インデックスは、キャッシュ内アドレスを生成するために用いられる。バイトセレクトは、ライン単位のデータ群の中から所望のデータを選択して取得するために用いられる。

【0019】

図4は、16ビットで表されたメインメモリアドレスを3つのフィールドに分解する場合の一例である。本実施の形態において1ラインは16バイトであるので、メインメモリアドレスの下位4ビットをバイトセレクトとして用いる。タグフィールドとインデックスフィールドの分け方は、領域管理部13から送られてくる領域のサイズに応じて変わる。例えば、指定された領域のサイズが256ラインの場合、メインメモリアドレスの上位4ビットをタグフィールド、続く8ビットをインデックスフィールドとする。これは、インデックスの8ビットで指定された領域の256ラインを表すためである。また、指定された領域のサイズが128ラインであれば、メインメモリアドレスの上位5ビットをタグフィールド、続く7ビットをインデックスフィールドとする。インデックスは、キャッシュ内アドレス生成部15に送られ、タグはhit/mis判定部16に送られる。

【0020】

キャッシュ内アドレス生成部15は、領域管理部13から送られてきたベースアドレスと、アクセスアドレス分解部14から送られてきたインデックスを加算して、キャッシュ内アドレスを生成する。生成したキャッシュ内アドレスは、hit/mis判定部16に送られる。

hit/mis判定部16は、キャッシュディレクトリを参照して、キャッシュ内アドレス生成部15から送られてきたキャッシュ内アドレスとアクセスアドレス分解部14から送られてきたタグが対応して記憶されているか否かを判定する。判定が肯定的である場合、キャッシュメモリ制御装置1は、該当キャッシュ内アドレスが示すキャッシュメモリ上の位置に格納されているラインから、バイトセレクトを用いて、所望のデータを読み出して、マイクロプロセッサ3に送出する。判定が否定的であれば、キャッシュメモリ制御装置1は、主記憶制御装置4に所望のデータを含むライン単位のデータ群を読み出させ、生成したキャッ

シュ内アドレスが示す位置にデータ群を格納する。そして、キャッシュディレクトリに生成したキャッシュ内アドレスとタグを対応付けて記憶する。

【0021】

<動作>

ここで、マイクロプロセッサ3からメインメモリ5に格納されているデータへのリードアクセスにおいて、キャッシュメモリ制御装置1が行うリードアクセス処理について、フロー図を用いて説明する。

図5は、キャッシュメモリ制御装置1が行うリードアクセス処理を表すフロー図である。まず、キャッシュメモリ制御装置1のタスクID入力レジスタ11において、マイクロプロセッサ3からタスクIDを受け付ける（ステップS1）。受け付けたタスクIDにより、領域指定部12において、領域を指定する（ステップS2）。指定した領域のサイズを領域管理表から参照して特定し、アクセスアドレス分解部14において、特定したサイズに応じて、マイクロプロセッサ3より受け付けたメインメモリアドレスをタグ、インデックス、バイトセレクトに分解する。続いて、キャッシュアドレス生成部15において、インデックスと、ベースアドレスを加算することでキャッシュ内アドレスを生成する（ステップS3）。

【0022】

hit/mis判定部16において、キャッシュディレクトリを参照し、生成されたキャッシュ内アドレスと、アドレス分解により生成されたタグが対応して記憶されているか否かを判定する（ステップS4）。判定が肯定的である場合（ステップS4：Yes）、キャッシュ内アドレスが示すキャッシュメモリ上の位置に格納されているラインから、バイトセレクトを用いて、所望のデータを読み出し（ステップS5）、ステップS8に進む。

【0023】

一方、判定が否定的である場合（ステップS4：No）、キャッシュメモリ制御装置1は、主記憶制御装置4に、所望のデータを含むライン単位のデータ群を読み出させる（ステップS6）。続いて、読み出させたデータ群を生成したキャッシュ内アドレスが示すキャッシュメモリ上の位置に格納し（ステップS7）、

生成したキャッシュ内アドレスとタグを対応付けてキャッシュディレクトリに記憶して、ステップ S 8 に進む。

【0024】

ステップ 8 において、キャッシュメモリ制御装置 1 は、マイクロプロセッサ 3 にデータを送出して、リードアクセス処理を終了する。

<補足>

なお、本発明は上記実施の形態に限定されないのはもちろんである。以下のような場合も本発明に含まれる。

(1) 実施の形態において、キャッシュメモリ制御装置 1 は、マイクロプロセッサ 3 からタスク ID を受け付けて領域指定を行っていたが、マルチプロセッサや、複数のプロセッサからプロセッサ ID を受け付けて領域指定を行ってもよい。

(2) 実施の形態において、各領域と各タスクとの対応関係は 1 対 1 であるが、複数のタスクが、同じメインメモリアドレスのデータをアクセスする確率が高い場合、1 つの領域と複数のタスクとを対応付けて設定してもよい。

(3) 実施の形態において、キャッシュメモリ制御装置 1 は、各領域と各タスクとの対応付け及び領域のサイズを、マイクロプロセッサ 3 からの設定命令（プログラム）を受けて設定したが、他の外部装置から設定命令を受けて設定してもよい。また、マイクロプロセッサ 3 又は他の外部装置から受けた設定命令により、動的に領域の生成及び消去が可能であるとしてもよい。

(4) 実施の形態において、hit/miss 判定部 16 は、領域を限定せずにキャッシュディレクトリを参照し、生成されたキャッシュ内アドレスと、アドレス分解により生成されたタグが対応して記憶されているか否かを判定していたが、キャッシュディレクトリに記載されているキャッシュ内アドレスのうち、指定された領域内のキャッシュ内アドレスのみを参照するようにしてもよい。

(5) 実施の形態において説明したコンピュータシステムに、MMU (memory management unit) が含まれていてもよい。MMU は、メモリのマッピングや保護を目的としたメモリ管理のためのハードウェアであり、物理的な記憶領域をブロック単位で管理し、それを論理的なメモリ空間にマッピングする機能と、プログラムによってアクセスできるメモリ領域を制限するメモ

り保護機能を備える。

(6) 本発明は、上記実施の形態に示すキャッシュメモリ制御方法であるとしてもよい。また、キャッシュメモリ制御方法を実現する制御プログラムであるとしてもよいし、前記制御プログラムからなるデジタル信号であるとしてもよい。

(7) 本発明は、前記制御プログラム又は前記デジタル信号をコンピュータ読み取り可能な記録媒体、例えば、フレキシブルディスク、ハードディスク、CD-ROM、MO、DVD、DVD-ROM、DVD-RAM、BD (Blu-ray Disc)、半導体メモリ等に記録したものとしてもよい。

(8) 本発明は、前記制御プログラム又は前記デジタル信号を、電気通信回線、無線又は有線通信回線、インターネットを代表とするネットワーク等を経由して伝送するものとしてもよい。

(9) 本発明は、前記制御プログラム又は前記デジタル信号を前記記録媒体に記録して移送することにより、又は前記制御プログラム又は前記デジタル信号を前記ネットワーク等を経由して移送することにより、独立した他のコンピュータシステムにより実施するものとしてもよい。

【0025】

【発明の効果】

以上説明したように、本発明に係るキャッシュメモリ制御装置は、マイクロプロセッサ、キャッシュメモリ及びメインメモリを含み、複数のタスク処理を行うコンピュータシステムにおいて、キャッシュメモリを制御するキャッシュメモリ制御装置であって、前記マイクロプロセッサから、実行中のタスクを識別するタスクIDを受け付けるタスクID受付手段と、前記マイクロプロセッサから、メインメモリ上のアドレスを受け付けるアドレス受付手段と、前記複数のタスクそれぞれに前記キャッシュメモリ内の別個のキャッシュメモリ領域を割り当てている割当手段と、前記アドレス受付手段により受け付けたアドレスのデータが前記キャッシュメモリに格納されていない場合、前記メインメモリから当該アドレスのデータを含むデータ群を取得し、当該データ群を前記タスクID受付手段により受け付けたタスクIDと対応するキャッシュメモリ領域に格納する格納手段とを備えることを特徴とする。

【 0 0 2 6 】

これにより、あるタスク実行による新たなキャッシングにより、他のタスク実行によってキャッシングされたデータ群が書き替えられないようにすることが可能となり、各タスクのリアルタイム性を保証することができる。

【図面の簡単な説明】

【図 1】 本発明に係るコンピュータシステムのブロック構成図である。

【図 2】 本発明に係るキャッシュメモリ制御装置のブロック構成図である。

。

【図 3】 領域管理表を示す図である。

【図 4】 1 6 ビットで表されたアドレスを 3 つのフィールドに分解する場合の一例図である。

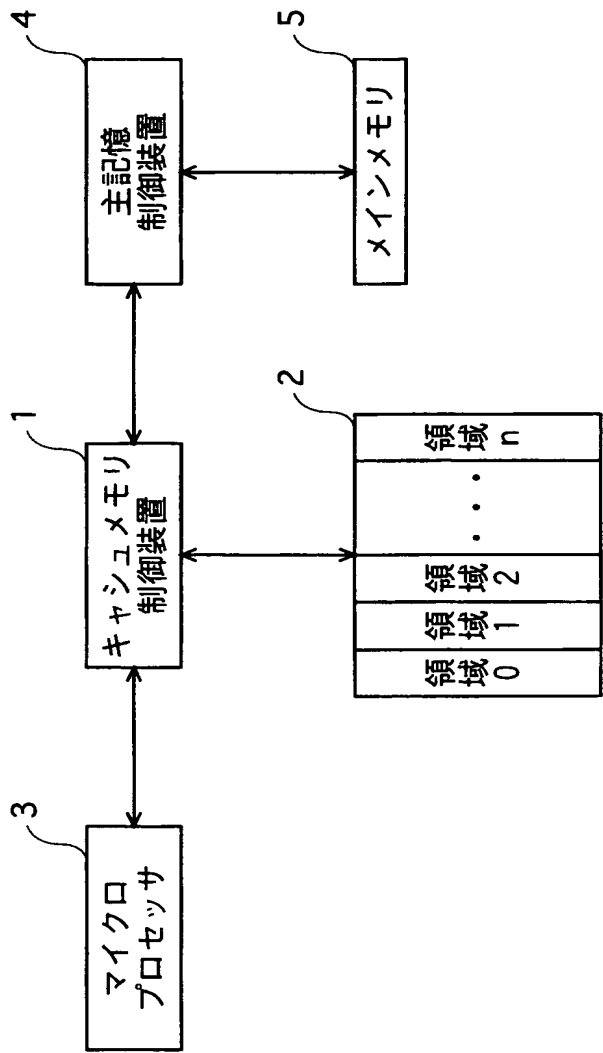
【図 5】 本発明に係るキャッシュメモリ制御装置が行うリードアクセス処理を示すフロー図である。

【符号の説明】

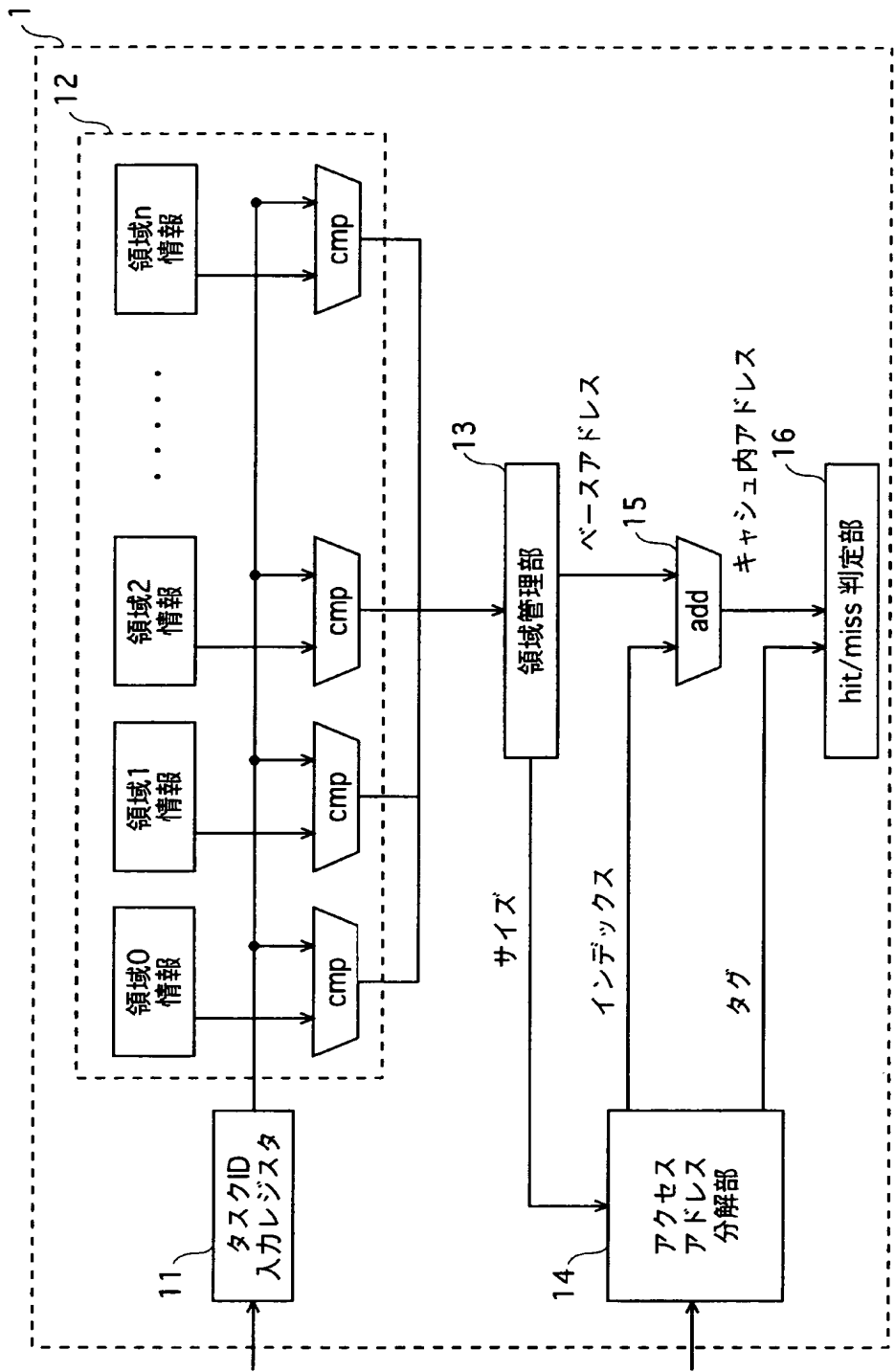
- 1 キャッシュメモリ制御装置
- 2 キャッシュメモリ
- 3 マイクロプロセッサ
- 4 主記憶制御装置
- 5 メインメモリ
- 1 1 タスク I D 入力レジスタ
- 1 2 領域指定部
- 1 3 領域管理部
- 1 4 アクセスアドレス分解部
- 1 5 キャッシュ内アドレス生成部
- 1 6 h i t / m i s s 判定部

【書類名】 図面

【図 1】



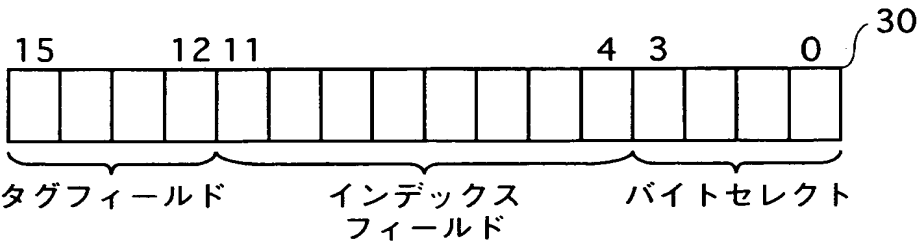
【図 2】



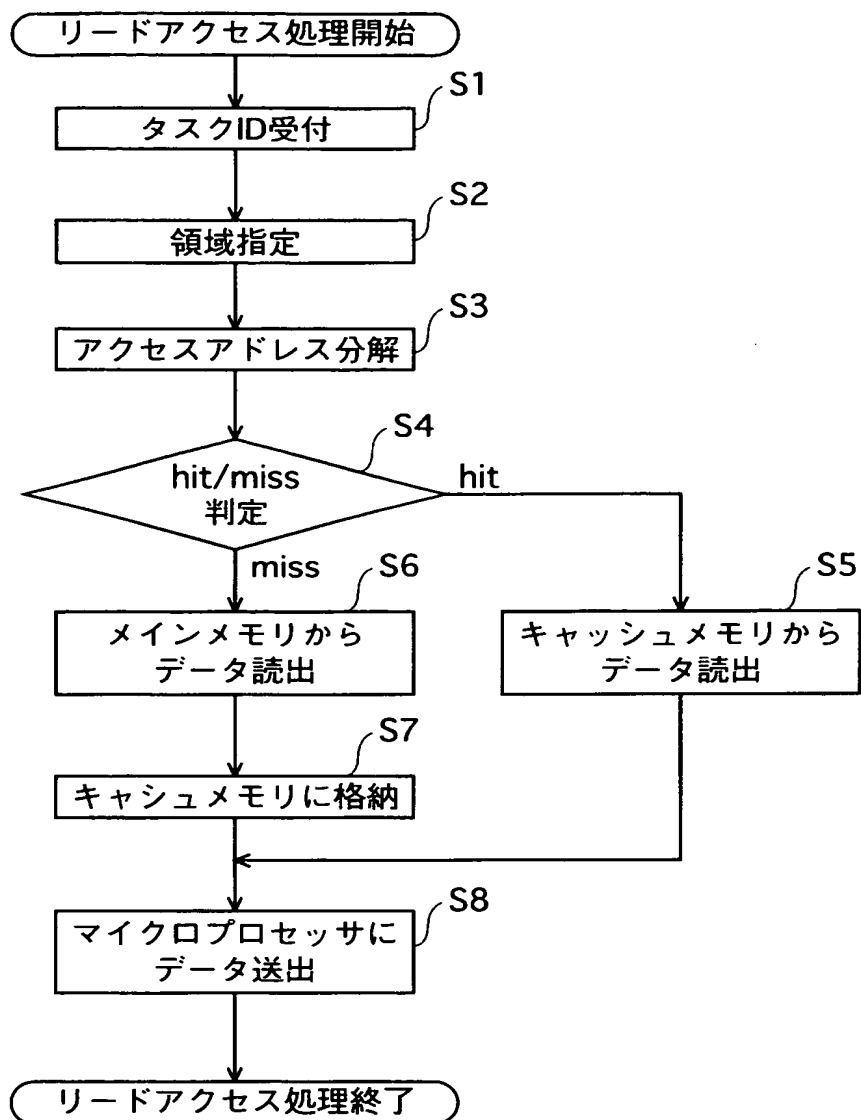
【図 3】

領域番号	ベースアドレス	サイズ
0	0	256
1	256	128
2	384	64
⋮	⋮	⋮
n	2304	128

【図 4】



【図 5】



【書類名】 要約書

【要約】

【課題】 シングルマイクロプロセッサによるマルチタスク処理や複数のマイクロプロセッサ又はマルチプロセッサによる複数のタスク処理において、各タスクのリアルタイム性を保証するべく、あるタスク実行による新たなキャッシングにより、他のタスク実行によってキャッシングされたデータ群が書き替えられないようにする。

【解決手段】 各タスクの実行前に、キャッシュメモリ内の別個のキャッシュメモリ領域をタスク毎に割り当てておき、マイクロプロセッサから受け付けたメインメモリアドレスのデータを含むデータ群がキャッシュメモリに格納されていない場合、メインメモリから当該アドレスのデータを含むデータ群を取得し、当該データ群をマイクロプロセッサから受け付けたタスク ID と対応するキャッシュメモリ領域に格納する。

【選択図】 図 2

特願 2 0 0 2 - 3 2 7 4 2 9

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 8 2 1]

1 . 変 更 年 月 日

1 9 9 0 年 8 月 2 8 日

[変 更 理 由]

新 規 登 録

住 所

大 阪 府 門 真 市 大 字 門 真 1 0 0 6 番 地

氏 名

松 下 電 器 産 業 株 式 会 社